Small Worlds and Rapid Mixing with a Little More Randomness on Random Geometric Graphs

Gunes Ercal

gunes@ku.edu
University of Kansas

Abstract. We theoretically and experimentally analyze the process of adding sparse random links to a random wireless networks modeled as a random geometric graph. While this process has been previously proposed, we are the first to (i) theoretically consider sparse new-wiring (with asymptotically less than constant fraction of nodes allowed very small constant new wired edges), (ii) prove bounds for any new-wiring process upon random geometric graphs, and (iii) consider the effect of such sparse new-wiring upon the spectral gap of the resultant normalized Laplacian. In particular, we consider the following models of adding new wired edges: Divide the normalized space into bins of length $k\frac{r}{2\sqrt{2}} \times k\frac{r}{2\sqrt{2}}$, given that the radius is on the order required to guarantee asymptotic connectivity. For each bin, choose a bin-leader. Let the G_1 new wiring be such that we form a random cubic graph amongst the bin-leaders and superimpose this upon the random geometric graph. Let the G_2 new wiring be such that we form a random 1-out graph amongst the bin-leaders and superimpose this upon the random geometric graph. We prove that the diameter for G_1 is $O(k + log(n))$ with high probability, and the diameter for G_2 is $O(klog(n))$ with high probability, both of which exponentially improve the $\Theta(\sqrt{\frac{n}{\log n}})$ diameter of the random geometric graph, thus also inducing small-world characteristics as the high clustering remains unchanged. Further, we theoretically motivate and experimentally demonstrate that the spectral gap for both G_1 and G_2 are significantly greater than that of the original random geometric graph. These results further motivate future hybrid networks and advances in the use of directional antennas.

Keywords: small world, wireless networks, spectral gap, theory

1   Introduction and Related Work

Ever since the first observation of "six degrees of separation" by Stanley Milgram [1], small-world phenomenon have been noted in numerous diverse network domains, from the World Wide Web to scientific co-author graphs [2]. The pleasant aspect of the small-world observations is that, despite the high clustering

characteristic of relationships with "locality", these various real world networks nonetheless also exhibit short average path lengths as well. This is surprising because purely localized graphs such as low dimensional lattices have very high average path lengths and diameter, whereas purely non-localized graphs such as random edge graph models of Erdos and Renyi [3] exhibit very low clustering coefficient. With intuition consolidating these two extremal graph types, the first theoretical and generative model of small world networks was proposed by Watts and Strogatz [4]: Start with a one dimensional k-lattice, and re-wire every edge to a new uniformly at random neighbor with a small constant probability. They showed that even for a very small but constant re-wiring probability, the resulting graph has small average path lengths while still retaining significant clustering.

Despite the prevalence of small world phenomenon in many real-world networks, wireless networks, in particular ad-hoc and sensor networks, do *not* exhibit the small average path lengths required of small-world networks despite the evident locality arising from the connectivity of geographically nearby nodes. Although taking a high enough broadcast radius r clearly can generate a completely connected graph of diameter one, this is a non-realistic scenario because energy and interference also grow with r. Rather, from a network design and optimization perspective, one must take the smallest reasonable radius from which routing is still guaranteed. To discuss such a radius in the first place, we must employ a formalization which is common in all theoretical work on wireless networks [5, 6], namely we fix the *random geometric graph* model of wireless networks. Given parameters n, the number of nodes, and r, the broadcast radius, the random geometric graph $\mathcal{G}(n,r)$ is formed by uniformly at random dispersing the n nodes into the unit square (which is a normalized view of the actual space in which the nodes reside), and then connecting any two nodes iff they are within distance r of each other. Note that due to the normalization of the space, r is naturally viewed as a function of n. Given such a model, it is a seminal result of Gupta and Kumar [6] that the *connectivity* property exhibits a sharp threshold for $\mathcal{G}(n,r)$ at critical radius $r_{con} = \sqrt{\frac{\log n}{\pi n}}$, which also corresponds to an average degree of $\log n$. As connectivity is a minimal requirement for routing, r_{con} is the reference point to take for analysis of $\mathcal{G}(n,r)$, and yet, as we shall see, such a radius still yields average path lengths of $\Theta(\sqrt{\frac{\pi n}{\log n}})$ with high probability.

This serves as a first motivation for the question: In the spirit of small-world generative models [4] that procured short average path lengths from a geographically defined lattice by adding random "long" edges, can we obtain significant reduction in path lengths by adding random "short cut" wired links to a wireless network? The first to ask this question in the wireless context was Ahmed Helmy [7] who experimentally observed that even using a small amount of wires (in comparison to network size n) that are of length at most a quarter of the physical diameter of the network yields significant average path lengths reduction. Another seminal work on this question is that of Cavalcanti et al. [8] which showed that introducing a fraction f of special nodes equipped with two radios,

one for short-range transmission and the other for long-range transmission, improves the connectivity of the network, where this property is seen to exhibit a sharp threshold dependent on both the fraction f and the radius r. Other work yet include an optimization approach with a specified sink, in which the placement of wired links is calculated to decrease average path lengths in the resulting topology [9]. The existing body of literature authored by practitioners in the field of wireless networks on inducing small-world characteristics (particularly shortened average path lengths) into wireless networks by either introducing wired links or nodes with special long range radios or directional antennas yields that such hybrid scenarios are eminently reasonable to consider for real networks. Nonetheless, no theoretical bounds have yet been proven for such processes of long-range link addition on random geometric graphs. We establish such bounds in this work for models of sparse random wired link additions, with the sparsity controlled by a parameter k.

In particular, we consider the following models of adding new wired edges: Divide the normalized space into bins of length $k\frac{r}{2\sqrt{2}} \times k\frac{r}{2\sqrt{2}}$, given that the radius is on the order required to guarantee asymptotic connectivity. For each bin, choose a bin-leader. Let the G_1 new wiring be such that we form a random cubic graph amongst the bin-leaders and superimpose this upon the random geometric graph. Let the G_2 new wiring be such that we form a random 1-out graph amongst the bin-leaders and superimpose this upon the random geometric graph. We prove that the diameter for G_1 is $O(k + log(n))$ with high probability, and the diameter for G_2 is $O(klog(n))$ with high probability, both of which exponentially improve the $\Theta(\sqrt{\frac{n}{\log n}})$ diameter of the random geometric graph, thus also inducing small-world characteristics as the high clustering remains unchanged. Our results on resulting average path lengths are also stable in comparison to using a constant fraction of wire lengths, as that in the work of Ahmed Helmy [7]. To see this note that, for example, using a maximum wire length of one-quarter the maximum distance can be simulated by subdividing the unit square into 16 parts and applying results to the parts separately, then combining into a maximum average path length that is still at most 16 of that within each part.

In addition to establishing bounds on resultant path lengths, we are the first to consider the resultant *spectral gap*, namely the difference between the first and second eigenvalues of the normalized adjacency matrix, in comparison to that of the random geometric graph. It is well established that the spectral gap is intrinsically related to the node expansion, edge expansion [10], connectivity, and random walk properties of the given graph [11, 12]. Although short average path lengths is necessary for a graph to exhibit optimal random walk sampling properties, it is far from sufficient (the barbell graph being a notable counterexample). As a strange omission, the small world literature thus far has primarily ignored spectral gap as a measure in their analyses and generative models despite the known expansion of random edge graph models.

Our motivation is as follows: Yet another limitation of random geometric graphs in comparison to random edge graph models that is especially problematic for oblivious routing, sampling, and gossiping applications[13–15] is that,

whereas sparse random regular graphs as well as random connected Erdos-Renyi graphs have a large spectral gap [16, 17], connected random geometric graphs $\mathcal{G}(n, \Theta(r_{con}))$ have asymptotically quickly diminishing spectral gap [18–20]. In general, additional edges need not improve the spectral gap of the resulting graph. However, the author notes that the *conductance argument* used to bound the spectral gap of random geometric graphs in [19], would no longer immediately yield bad bounds using the same counterexample of dividing down the middle if random edges are added to the graph with any sparseness so long as the number of such edges grows asymptotically in network size. Motivated by this observation, we experimentally calculate the graph spectra, however using the normalized Laplacian which is guaranteed to be symmetric and similarly is intrinsically related to random walk sampling properties [21], and observe that indeed there is an asymptotic improvement in the spectral gap of G_1 and G_2 in comparison to the starting random geometric graph.

Finally, we must note the work [22] of Abraham Flaxman, which is an excellent related work in which the spectra of randomly *perturbed* graphs have been considered in a generality that already encompasses major small-world models thus far. [22] demonstrates that, no matter what is the starting graph G_0, adding random 1-out edges at every node of G_0 will result in a graph with constant spectral gap (the best possible asymptotically). The work also presents a condition in which a random Erdos-Renyi graph superimposed upon the nodes of G_0 would yield good expansion, whereas without that condition the resulting graph may have poor expansion. Despite the apparent generality of the work in terms of the arbitrariness of the underlying graph that is considered, unfortunately the results do not generalize to situations in which not all nodes are involved in a wired linkage. Notably, the small world models thus far also require such a high probability of new random links. In contrast, in this work, we focus on adding sparse random wires. In particular, the fraction of nodes involved in a wired link will be no more than $O(\frac{1}{\log n})$, and in general shall be $O(\frac{1}{k^2 \log n})$, both of which are asymptotically diminishing fractions.

2 Theoretical Results

2.1 Random Geometric Graph Preliminaries

Our analytical results for modifications of random geometric graphs are based on certain "smooth lattice-like" properties that those graphs have. These properties include the uniformity of nodes distribution and the regularity of node degree. As introduced in [20], we utilize the notion of a *geo-dense* graph to characterize such properties, that is, a geometric graph (random or deterministic) with uniform node density across the unit square. It was shown that *random geometric graphs* are *geo-dense* and for radius $r_{\text{reg}} = \Theta(r_{\text{con}})$ all nodes have the same order degree[20]. We formally present the relevant results from [20] in this section, as well as the notion of *bins*, namely the equal size areas that partition the unit square. Such "bins" are the concrete link between lattices and random geometric graphs, essentially forming the lattice backbone of such graphs.

Formally, a **geometric graph** is a graph $G(n, r) = (V, E)$ with $n = |V|$ such that the nodes of V are embedded into the unit square with the property that $e = (u, v) \in E$ if and only if $d(u, v) \leq r$ (where $d(u, v)$ is the Euclidean distance between points u and v). In wireless networks, r naturally corresponds to the broadcast radius of each node. The following formalizes geo-denseness for geometric graphs:

Definition 21 [20] *Let $G(n, r(n))$ be a geometric graphs* [1]. *For a constant $\mu \geq 1$ we say that such a class is μ-**geo-dense** if every square bin of size $A \geq r^2/\mu$ (in the unit square) has $\Theta(nA)$ nodes.* [2]

The following states the almost *regularity* of geo-dense geometric graphs[20]:

Lemma 22 *Let $G(n, r)$ be a 2-geo-dense geometric graph with V the set of nodes and E the set of edges. Let $\delta(v)$ denote the degree (i.e number of neighbors) of $v \in V$. Then: (i) $\forall v \in V$ $\delta(v) = \Theta(nr^2)$ and (ii) $m = |E| = \Theta(n^2r^2)$*

Recall that the critical radius for connectivity r_{con} is s.t $\pi r_{\text{con}}^2 = \frac{\log n}{n}$. The following is the relevant lemma that states that *random* geometric graphs with radius at least on the order of that required for connectivity are indeed geo-dense[20]:

Lemma 23 (Geo-density of $\mathcal{G}(n, r)$) *For constants $c > 1$ and $\mu \geq 2$, if $r^2 = \frac{c\mu \log n}{n}$ then w.h.p. $\mathcal{G}(n, r)$ is μ-geo-dense. That is, w.h.p. (i) any bin area of size r^2/μ in $\mathcal{G}(n, r)$ has $\Theta(\log n)$ nodes, and (ii) $\forall v \in \mathcal{G}(n, r)$, $\delta(v) = \Theta(nr^2)$ and $m = |E| = \Theta(n^2r^2)$. Further, note that increasing the radius r can only smoothen the distribution further, maintaining regularity.*

Given that geo-denseness of connected random geometric graphs is established, we wish to utilize the "binning" directly for its lattice-like properties. As such, we shall introduce the notion of a *lattice skeleton* for geo-dense geometric graphs, including random geometric graphs above connectivity:

Definition 24 *Let $G(n, r(n))$ be a μ-geo-dense geometric graph for some constant $\mu \geq 1$. Define the μ-lattice skeleton for $G(n, r)$ to be $LS(G(n, r)) = (L, B_{i,j})$ such that $L = (B, E_B)$ is the $\frac{\sqrt{\mu}}{r} \times \frac{\sqrt{\mu}}{r}$ two-dimensional lattice on $\frac{\mu}{r^2}$ bin-points, where each vertex set $B_{i,j}$ represents the set of nodes of $G(n, r)$ that lie in lattice location $\{i, j\}$ of L. Further, for a node $v \in G(n, r)$ with Cartesian coordinates $(x, y) \in [0, 1]^2$, denote by $B(v)$ the lattice-bin containing v, namely the bin $B_{i,j}$ such that $i = ceiling(\frac{x\sqrt{\mu}}{r})$ and $j = ceiling(\frac{y\sqrt{\mu}}{r})$.*

What is directly clear by geo-denseness is that there is not much variance in the sizes of the bins:

---
[1] either random or deterministic
[2] Note that if a geometric graph is μ_1-geo-dense then it is also μ_2-geo-dense for any $\mu_2 < \mu_1$. That is, if a distribution is smooth for some granularity x it can only be smoother for a coarser granularity y.

Remark 25 *Let $LS(G(n,r)) = (L, B_{i,j})$ be the μ-lattice skeleton of a μ-geo-dense geometric graph $G(n,r)$. Then, $\forall i, j, |B_{i,j}| = \Theta(nr^2)$.*

Further, utilizing the choice of $\mu \geq 5$, we may make the stronger statement that the connectivity of the lattice is inherited in the nodes of the overall graph. The justification is simply that r becomes the length of the diagonal connecting the farthest points of adjacent bins, and we formalize combining with Remark 25 and Lemma 23:

Lemma 26 *Let $LS(G(n,r)) = (L, B_{i,j})$ be the 5-lattice skeleton of a 5-geo-dense geometric graph $G(n,r)$. Then $\forall i, j, k, l$ if $((i,j),(k,l)) \in L$ then $\forall v \in B_{i,j}, u \in B_{k,l}, (v,u) \in G(n,r)$*
In particular, for $c > 1$ if $r^2 \geq \frac{5c}{\pi} r_{con}^2$ then the 5-lattice skeleton $LS(\mathcal{G}(n,r)) = (L, B_{i,j})$ of random geometric graph $\mathcal{G}(n,r)$ satisfies the following:
(i) $\forall i, j |B_{i,j}| = \Theta(nr^2)$ w.h.p.,
(ii) $\forall i, j$ if $v, u \in B_{i,j}$ then $(v,u) \in \mathcal{G}(n,r)$,
(iii) $\forall i, j, k, l$ if link $((i,j),(k,l)) \in L$ then $\forall v \in B_{i,j}, \forall u \in B_{k,l}, (v,u) \in \mathcal{G}(n,r)$, and
(iv) $\forall i, j, k, l$ if $\min\{|i-k|, |j-l|\} \geq 4$ then $\forall v \in B_{i,j}, \forall u \in B_{k,l}, (v,u) \notin \mathcal{G}(n,r)$

From (ii) it is clear that each bin $B_{i,j}$ forms a clique (namely all pairs of nodes within are connected directly by length one paths). From (iii) it follows that a path in the lattice L yields a path in the graph $\mathcal{G}(n,r)$ as well, while (iv) bounds the converse situation in that nodes that lie in bins at least 4 lattice-hops away cannot be directly connected in the graph $\mathcal{G}(n,r)$ either. In particular, (iii) and (iv) yield that pairwise distances between points in the graph $\mathcal{G}(n,r)$ inherit the shortest paths (Manhattan) distances in the corresponding lattice-bins of the lattice-skeleton, up to constant factors. We formalize with the following corollary:

Corollary 27 *For $c > 1$ if $r^2 \geq \frac{5c}{\pi} r_{con}^2$ then the 5-lattice skeleton $LS(\mathcal{G}(n,r)) = (L, B_{i,j})$ of random geometric graph $\mathcal{G}(n,r)$ satisfies the following w.h.p.: $\forall u, v \in \mathcal{G}(n,r), dist_{\mathcal{G}(n,r)}(u,v) = \Theta(dist_L(B(u), B(v)) + 1)$ where the function $dist_G$ indicates shortest paths distances in graph G.*

Having established that connectivity and distances for $\mathcal{G}(n,r)$ with radius at least a small constant times r_{con} roughly preserve connectivity and distances in the $\frac{\sqrt{\mu}}{r} \times \frac{\sqrt{\mu}}{r}$ lattice skeleton, let us then consider the number of lattice-nodes $N_{d,L}(v)$ that are at lattice distance exactly d away from v in the lattice L: Clearly, $N_{d,L}(v)$ grows linearly in d by a simple induction on upper and lower bounds. And, the maximum distance to consider is $d = \Theta(\frac{\mu}{r^2})$. Moreover, due to the smooth distribution of random geometric graph nodes in the lattice bins, we must have that the *fraction* $f_{d,L}(B(v))$ of lattice bins at lattice-distance exactly d away from $B(v)$ must be on the same order as the *fraction* of random geometric graph nodes $f_{d,\mathcal{G}(n,r)}(v)$ at hop-distance exactly d away from v. Thus,

$$f_{d,\mathcal{G}(n,r)}(v) = \Theta(f_{d,L}(B(v))) = \Theta(\frac{N_{d,L}(v)}{|L|}) = \Theta(\frac{N_{d,L}(v)r^2}{\mu}) = \Theta(dr^2)$$

Such fractions represent the probability that a node is at distance d away from

a given node v. Thus, we may calculate the average path length APL which is the expectation of that probability distribution on the very function d itself:

$$APL(\mathcal{G}(n,r)) = \sum_{d=1}^{D(\mathcal{G}(n,r))} df_{d,\mathcal{G}(n,r)}(v) = \Theta(\sum_{d=1}^{2\frac{\sqrt{\mu}}{r}} df_{d,L}(B(v))) = \Theta(\sum_{d=1}^{2\frac{\sqrt{\mu}}{r}} d^2 r^2) = \Theta(r^2 \sum_{d=1}^{2\frac{\sqrt{\mu}}{r}} d^2) = \Theta(r^2 \frac{1}{r^3}) = \Theta(\frac{1}{r})$$

Thus, the average path length for random geometric graphs above the connectivity threshold is the same as the order of the diameter (maximum shortest path lengths) for such graphs, which is $\Theta(\frac{1}{r})$. While the dependence on the radius r in that term may seem optimistic at first, noting that r should be kept as low as possible to reduce energy overhead and interference of the ad-hoc network represented, a realistic constraint on r becomes $r = \Theta(r_{con}) = \Theta(\sqrt{\frac{\log n}{n}})$, namely that achieving degree $\Theta(\log n)$. Thus, APL of reasonable random geometric graphs (of minimal radius guaranteeing connectivity) scales quite badly as $\Theta(\sqrt{\frac{n}{\log n}})$.

2.2 Adding Random Edges

Prior to considering the business of adding random edges to a given initial graph $G_0 = (V_0, E_0)$, note that the set of additional edges E_R and the existing nodes connected by them $V_R \subset V$ forms a graph G_R such that the resulting graph $G = (V, E) = G_0 + G_R$ has $V = V_0 \geq V_R$ and $E = E_0 + E_R$. That is, it is also convenient to view the additional random edges as a new graph G_R superimposed upon the original graph G_0.

Given such a characterization, let us be given a 5-geo-dense geometric graph $G_0 = (V_0, E_0) = G(n, r)$ with 5-lattice-skeleton $(L, B_{i,j})$. In particular, note from Lemma 23 that results apply to any $G_0 = \mathcal{G}(n, \Omega(r_{con}))$. Given parameter $k \geq 1$, let vertex set $V_R(k)$ be generated as follows: For any $i, j \leq \frac{\sqrt{5}}{\sqrt{k}r}$ pick a node $v_{i,j}$ uniformly at random from the nodes in the set of bins $\mathcal{B}_{i,j} = \cup_{ki \leq i' \leq (k+1)i, kj \leq j' \leq (k+1)j} B_{i',j'}$[3], and set $v_{i,j} \in V_R(k)$. For the case $k = 0$, let $V_R(0) = V_0$. We now define the various types of random edge sets $E_{R,i}$ for graphs $G_{R,i}(k) = (V_R(k), E_{R,i})$ whose superimpositions upon G_0 we shall consider in this work: Let $E_{R,1}$ be generated as follows: For every node $v \in V_R(k)$ pick 3 neighbors in $V_R(k)$ uniformly at random, discarding situations in which any node has degree greater than 3. Thus, the resulting graph $G_{R,1}(k) = (V_R(k), E_{R,1})$ is the random 3-regular Erdos-Renyi graph defined on vertex set $V_R(k)$. Let $E_{R,2}$ be generated as follows: For every node $v \in V_R(k)$ pick 1 neighbor in V_R uniformly at random. Thus, the resulting graph $G_{R,2}(k) = (V_R(k), E_{R,2})$ is the random 1-out graph defined on vertex set $V_R(k)$.

Similarly to the above, let us define the resulting graphs as follows: Let $G_1 = G_0 + G_{R,1}$. Let $G_2 = G_0 + G_{R,2}$ Essentially, k controls the frequency of

---

[3] This bin union is simply the formalization of a single contiguous bin $k \times k$ as large as the original bins.

special nodes which shall serve as wired link stations. For $k = \Theta(1)$, the frequency is in line exactly with the bins, and thus the occurence of such wired link stations is 1 in every $\Theta(nr^2)$. For $r = \Theta(r_{con})$ that frequency becomes $\Theta(\frac{1}{\log n})$, and for larger broadcast radius it is sparser:

Remark 28 *Given* $k \geq 1$*, the frequency of wired link stations* $\frac{|V_R(k)|}{|V_0|}$ *is* $\Theta(\frac{1}{k^2nr^2})$*. Namely, the total number of such stations is* $|V_R(k)| = \Theta(\frac{1}{k^2r^2})$*.*[4]

Before proceeding to prove results on average path lengths for $G_1 = G_0 + G_{R,i}$, we note that the manner in which $V_R(k)$ is generated can be simulated approximately by simply choosing a total of logarithmically more wired link stations uniformly at randomly from the original set V_0. This too follows from "coupon collection":

Remark 29 *For any* k*, if every* $v \in V_0$ *is chosen to be a wired link station with probability* $\Theta(\frac{\log(\frac{1}{k^2r^2})}{nk^2r^2})$*, then, with high probability, for every* k^2*-bin* $\mathcal{B}_{i,j} = \cup_{ki \leq i' \leq (k+1)i, kj \leq j' \leq (k+1)j} B_{i',j'}$ *there exists a vertex* $v' \in \mathcal{B}_{i,j}$ *such that* v' *is a wired link station. Moreover, all of the vertices in any* k^2 *bin are almost-equiprobable and almost-independent whp.*

The last is simply stating that there is not much lost from the uniformly at random choice of a finite set of nodes versus the corresponding Poisson process.

Now, we note that the maximum distance of any node in a k^2 bin to the corresponding wired link station in the k^2 bin is simply bounded by the hop-diameter of the k^2 bin:

Remark 210 *Every node is within* $\Theta(k)$ *hops of the wired link station in its* k^2*-bin since the* k^2*-bin simply stretches the Manhattan distances of the original 5-lattice-skeleton by* k*.*

This remark shall prove relevant in relating inter-node distances in the graph $G = G_0 + G_R$ to inter-node distances in G_R.

2.3 APL and Diameter bounds for G_1

Thus, we now proceed concerning inter-node distances for the first model $G_{R,1}$:

Remark 211 *It is well-known that the random 3-regular graph* $G_{R,1}(k)$ *is an expander with high probability [10, 16]. Therefore,* $G_{R,1}(k)$ *also exhibits diameter and average path lengths asymptotically at most logarithmic in its vertex set* $|V_R(k)|$*, with high probability.*

Combining Remarks 28, 210, and 211, we obtain our first bounds on the resulting average and worst-case path lengths:

Theorem 212 *The diameter and average path lengths for* $G_1(k)$ *are as follows:* $APL(G_1(k)) = O(Diam(G_1(k))) = O(\min\{2k + \log(\frac{1}{k^2r^2}), \frac{1}{r}\}) = O(\min\{k + \log(n), \frac{1}{r}\})$*.*

---

[4] Note that clearly k cannot exceed $\Theta(\frac{1}{r})$.

Proof. Let $s, t \in V$ be arbitrary. s is within k hops to a node $v \in V_R(k)$, and t is within k hops to a node $w \in V_R(k)$. Moreover, $dist(v, w) = O(\log(|V_R(k)|))$, making the maximum distance at most $2k + \log(|V_R(k)|)$. If r happens to be large, making paths in the original graph (without taking short-cuts in $E_{R,1}$ more convenient, then the maximum distance is the diameter of G_0 which is $\Theta(\frac{1}{r})$.

For *any* broadcast radius r, it is clear that when $k = \Theta(\frac{1}{r})$, which is the maximum allowable k and corresponds to the situation of placing a constant number $\Theta(1)$ wired-stations, there is no asymptotic improvement in diameters and average path lengths. However, for any other intermediate k and the lowest reasonable broadcast radius $r = \Theta(r_{con})$, the following may be noted:

Corollary 213 *For $k = O(\log(n))$ and $G_0 = \mathcal{G}(n, \Theta(r_{con}))$, $APL(G_1(k)) = O(Diam(G_1(k))) = O(\log(n))$. On the other hand, for intermediate k such that $k = o(\frac{n}{\log(n)})$ but $k = \omega(\log(n))$, we still obtain asymptotic improvement upon the hop-lengths in G_0 as $APL(G_1(k)) = O(Diam(G_1(k))) = O(k) = o(Diam(G_0))$.*

2.4 APL and Diameter bounds for G_2

The case for the second model G_2 is not quite as straightforward as that for the first model due to the fact that the random graph $G_{R,2}(k)$ has a positive probability of being disconnected. So, what can we say? It turns out that a lot can be said the moment that $G_{R,2}(k)$ is superimposed directly upon the nodes of *any* connected graph G_{init}:

Theorem 214 *[Expansion of $G_{connected} + R_{1-out}$][22] Let G_0 be any connected graph. Then $G = G_0 + R_{1-out}$, which is the graph formed by adding random 1-out edges to every vertex of G_0, is an expander with high probability. In particular, the diameter of G is logarithmic in the vertex set.*

Despite the similarities to the situation for G_1, there is the technical issue that the vertex set for $G_{R,2}(k)$ is not identical to the vertex set for G_0 but rather asymptotically sparser than such. Therefore, we must understand precisely which graph is an expander, and what that yields for the graph G_2. In particular, we need to extract a connected graph $G_{base} = (V_{base}, E_{base})$ that is both a relevant function of the original graph G_0 and such that there is a one-to-one meaningful correspondence between the vertex set V_{base} and $V_R(k)$ to allow application of Theorem 214.

Visually, if we could just contract the lattice-sub-skeleton formed from the k^2-bins into just a $|V_R(k)|$ node lattice, and preserve the meaning of such a contraction for paths in the original graph G_0, then we could apply Theorem 214 to our contraction, and then reverse the contraction. We may lose the property of constant expansion upon reversing our contraction for some k, but we will still preserve a bound on path lengths. First, let us state the result, then the proof along the idea above:

Theorem 215 *The diameter and average path lengths for $G_2(k)$ are as follows:*
$APL(G_2(k)) = O(Diam(G_2(k))) = O(\min\{k\log(|V_R(k)|), \frac{1}{r}\}) = O(k\log(\frac{1}{r}))$.

Prior to the theorem proof, let us note the immediate corollary for low broadcast radius, which again notes an exponential improvement in diameter and average path lengths in for poly-logarithmic k:

Corollary 216 *For* $k = \Theta(1)$ *and* $G_0 = \mathcal{G}(n, \Theta(r_{con}))$, $APL(G_2(k)) = O(Diam(G_2(k))) = O(\log(n))$. *On the other hand, for intermediate k such that* $k = o(\frac{n}{\log^2(n)})$ *but* $k = \omega(1)$, *we still obtain asymptotic improvement upon the hop-lengths in G_0 as* $APL(G_2(k)) = O(Diam(G_2(k))) = O(k\log n) = o(\frac{n}{\log(n)}) = o(Diam(G_0))$.

Now to the theorem proof:

Proof (Proof of Theorem 215). Consider the graph of node-contractions $G_{contract} = (V_{contract}, E_{contract})$ such that each vertex $v_{i,j} \in V_{contract}$ is precisely the *set* of vertices in k^2 bin $\mathcal{B}_{i,j}$ and the edge set $E_{contract}$ is the two-dimensional lattice appropriately defined on $V_{contract}$. We know from Theorem 214 that upon adding random 1-out edges from every $v_{i,j} \in V_{contract}$ we obtain a graph $G' = G_{contract} + R_{1-out}$ with diameter logarithmic in $V_{contract}$. What does this mean?

Consider any two nodes $s, t \in G_0$, say that $s \in \mathcal{B}_{i,j}, t \in \mathcal{B}_{i',j'}$. Let $l = < v_1, v_2, v_3, \ldots, v_{p-1}, v_p >$ be the shortest path between $v_{i,j}$ and $v_{i',j'}$ in $G_{contract}$ (with $v_1 = v_{i,j}, v_p = v_{i',j'}$). First of all, there is an obvious one-to-one correspondence between the vertex sets $V_{contract}$ and $V_R(k)$ which does not change the incidence of the random short-cut edges. The only issue occurs when paths in the contracted graph include a lattice edge, and in that case there is a factor $O(k)$ blow-up in the hop number for the original graph. Thus, for any path l in $G_{contract}$ we may inductively construct the following path l' in $G_{R,2}$ by a sequence of valid sub-path replacements in the place of contracted nodes and the edges between them until no contracted nodes and no contracted edges remain (meaning all nodes are actual vertices of our original graph):

- Base: In accordance with Remark 210, replace v_1 in l with the $O(k)$ length shortest-path from s to $v_{i,j}$ in G_0. Similarly, replace v_p in l with the $O(k)$ length shortest-path from $v_{i',j'}$ to t in G_0.

- Inductive case I: For $1 < t < p - 1$, let v_t and v_{t+1} be adjacent nodes in the path l that also remain thus far in our path construction. Uniquely define $v_{x,y} \in v_t, v_{x',y'} \in v_{t+1}$ for $v_{x,y}, v_{x',y'} \in V_R(k)$. Either edge (v_t, v_{t+1}) is a lattice edge or a random edge. If it is a random edge, then replace (v_t, v_{t+1}) of l with the valid wired edge $(v_{x,y}, v_{x',y'})$. Otherwise, if it is a lattice-edge, then there exists a $O(k)$ length shortest-path P between $v_{x,y}$ and $v_{x',y'}$ in G_0 from Lemma 26 and Remark 210. Thus, in that case, replace (v_t, v_{t+1}) with path P between $v_{x,y}$ and $v_{x',y'}$.

– Inductive case II: Let node $v_{x,y}$ be the node following initial wired-station $v_{i,j}$ in the construction thus far. Due to the construction, it must be that $v_{i,j} \in v_1, v_{x,y} \in v_2$. Therefore, similarly to Inductive case I, either there exists a random wired edge between $v_{i,j}$ and $v_{x,y}$, or there is a length $O(k)$ path between the two in G_0: Replace accordingly. Operate similarly in making the valid replacement between the node $v_{x',y'}$ preceding final wired station $v_{i'j'}$.

Clearly, the above construction is valid, replacing contracted edges with valid sub-paths at each step. Moreover, at every step, the replacement path is at most a $O(k)$ blow up of a contracted edge. Therefore, the constructed valid path in $G_{R,2}$ is also at most a $O(k)$ multiple of the diameter of $G_{contract}$, which we know by Theorem 214 to be logarithmic in $|V_R(k)|$, finalizing the proof.

3   Experimental Results
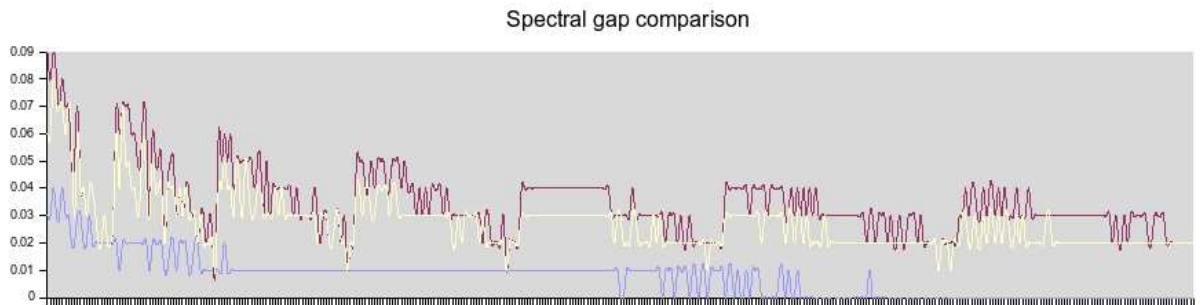
Fig. 1. Spectral Gap comparison for G_1 (Red), G_2 (Yellow), and $\mathcal{G}(n,r)$ (Blue) for network size from 100 to 1620 nodes, radius r_{con}, and k parameter of 2

Experiments were conducted for networks of 100 to 1620 nodes. As consistent with the model of the theoretical section, the parameter k was chosen to be 2, and the radius was chosen to be r_{con} exactly. Disconnected $\mathcal{G}(n,r)$ were discarded from consideration. The results can be seen in Figure 1, where the Y-axis is the spectral gap of the normalized Laplacian. Notably, the spectral gap for the random geometric graph approaches zero quickly, whereas the spectra gap for G_1 and G_2 appear to diminish concavely after 500 nodes, possibly even approaching a small constant lower bound. Moreover, note that the number of wired-nodes in comparison to the network size n for n values of 100, 300, 800, 1000, 1300, and 1620 are respectively as follows: 36, 64, 196, 256, 256, 324. The fraction of wired nodes for the network size of 1620 was just $\frac{1}{5}$.

4 Conclusion and Future Work

We have considered the "small world" effects of adding sparse random edges to an asymptotically diminishing frequency of nodes taken as wired link stations imposed upon a wireless network. In contrast to prior small-worlds work, we care in particular for very sparse link additions and control the sparseness of wired link stations by a parameter k. We establish theoretical proof that under two reasonable models of link additions (G_1 of superimposing a random cubic graph and G_2 of superimposing a random $1 - out$ graph) upon the random geometric graph model of wireless networks an exponential reduction of diameter and average path lengths is obtained even when the wired link stations occur with an inverse poly-logarithmic frequency. We also establish theoretical bounds on the diameter and average path lengths reduction as a general function of k (see Theorems 212 and 215).

As a related but separate question from the issue of average path lengths, we investigated the improvement in *spectral gap* as well under such models of sparse link additions. The spectral gap, which can refer to either the discrepancy between the first and second eigenvalues of the normalized adjacency matrix representing the graph or to the related measure of the second eigenvalue of the normalized Laplacian representing the graph in the cases of irregular graphs, is known to govern the rate at which the Markov chain corresponding to a natural random walk on the graph converges to the stationary distribution [11, 12]: The larger the spectral gap, the faster the convergence to stationarity (namely, the faster the *mixing rate*). It is intuitively clear that the faster the convergence to stationarity for such an oblivious process as a random walk, the greater the intrinsic load-balancing properties of the graph. This quantity of mixing rate is further intricately related to connectivity and resilience properties of the underlying graph measured as the combinatorial quantity of *expansion*. Expansion directly relates to resilience against edge failures, and when the graph is close to regular it also relates to resilience against node failures. In fact one characterization of an *expander graph family* is directly based on spectral gap alone: For a family of graphs parametrized by the size of the vertex set n, the graph family is an *expander graph family* iff the spectral gap is bounded away from zero by a constant, as n approaches ∞. Expander graphs represent extrema for the expansion property. While short average path lengths and short diameter are necessary for graphs to be expanders, these are far from sufficient. Thus, in investigating the improvement in expansion upon adding sparse random edges to a random geometric graph (which by itself is known to have horrible expansion [18–20]), we experimentally measured the spectral gap for the normalized Laplacian. We found that whereas the spectral gap for the pure random geometric graph falls rapidly to zero after some n (around 900), the spectral gap for G_1 and G_2 do not fall rapidly to zero but stay bounded away from zero by 0.03 for the duration of the experiment til $n = 1620$. These results are for the sparseness parameter $k = 2$. In either case, a definite discrepancy between the pure wireless spectral gap and the additional link models is exhibited.

Future work concerns establishing theoretical bounds on the spectral gap for G_1 and G_2 as a function of the parameter k. In particular: Is there a k for which G_1 or G_2 is an expander graph family? If so, does the expansion property exhibit a *sharp threshold* based on k? In addition to theoretical bounds, future work includes experimental measures of oblivious routing efficiency and load balancing quality based on random walk simulations of the respective graph families. And, finally, towards answering the question of improvement in resilience for the respective graph families, simulations of non-oblivious routing algorithms in the presence of node and edge failures may also be performed.

References

1. Milgram, S.: The small world problem. Psychology today **2**(1) (1967) 60–67
2. Newman, M., Barabási, A.L., Watts, D.J.: The Structure and Dynamics of Networks. Princeton University Press (2006)
3. Bollobás, B.: Random Graphs. Academic Press, Orlando, FL (1985)
4. Watts, D., Strogatz, S.: Collective dynamics of small-world networks. Nature **393**(6684) (1998) 440–442
5. Penrose, M.D.: Random Geometric Graphs. Volume 5 of Oxford Studies in Probability. Oxford University Press (May 2003)
6. Gupta, P., Kumar, P.R.: Critical power for asymptotic connectivity in wireless networks. In Stochastic Analysis, Control, Optimization and Applications: A Volume in Honor of W. H. Fleming (1998) 547–566
7. Helmy, A.: Small worlds in wireless networks. IEEE Communications Letters **7**(10) (2003) 490–492
8. Cavalcanti, D., Agrawal, D., Kelner, J., Sadok, D.: Exploiting the small-world effect to increase connectivity in wireless ad hoc networks. Telecommunications and Networking-ICT 2004 (2004) 388–393
9. Ye, X., Xu, L., Lin, L.: Small-world model based topology optimization in wireless sensor network. In: Proceedings of the 2008 International Symposium on Information Science and Engieering - Volume 01, Washington, DC, USA, IEEE Computer Society (2008) 102–106
10. Alon, N.: Eigenvalues and expanders. Combinatorica **6**(2) (1986) 83–96
11. Sinclair, A.: Improved bounds for mixing rates of markov chains and multicommodity flow. Combinatorics, Probability and Computing **1** (1992) 351–370
12. Jerrum, M., Sinclair, A.: The markov chain monte carlo method: an approach to approximate counting and integration. In: Approximations for NP-hard Problems, Dorit Hochbaum ed. PWS Publishing, Boston, MA (1997) 482–520
13. Kempe, D., Dobra, A., Gehrke, J.: Gossip-based computation of aggregate information. In: Proc. of the 44th Annual IEEE Symposium on Foundations of Computer Science. (2003) 482–491
14. Servetto, S.D., Barrenechea, G.: Constrained random walks on random graphs: routing algorithms for large scale wireless sensor networks. In: Proc. of the 1st Int. workshop on Wireless sensor networks and applications. (2002) 12–21
15. Gkantsidis, C., Mihail, M., Saberi, A.: Random walks in peer-to-peer networks. In: in Proc. 23 Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM). (2004)

16. Friedman, J., Kahn, J., Szemerédi, E.: On the second eigenvalue of random regular graphs. In: Proceedings of the twenty-first annual ACM symposium on Theory of computing. STOC '89, New York, NY, USA, ACM (1989) 587–598
17. Füredi, Z., Komlós, J.: The eigenvalues of random symmetric matrices. Combinatorica **1** (1981) 233–241
18. Boyd, S., Ghosh, A., Prabhakar, B., Shah, D.: Randomized gossip algorithms. IEEE/ACM Trans. Netw. **14**(SI) (2006) 2508–2530
19. Avin, C., Ercal, G.: Bounds on the mixing time and partial cover of ad-hoc and sensor networks. In: Proceedings of the 2nd European Workshop on Wireless Sensor Networks (EWSN 2005). (2005) 1–12
20. Avin, C., Ercal, G.: On the cover time and mixing time of random geometric graphs. Theor. Comput. Sci. **380**(1-2) (2007) 2–22
21. Chung, F.: Spectral Graph Theory. American Mathematical Society (1997)
22. Flaxman, A.D.: Expansion and lack thereof in randomly perturbed graphs. Internet Mathematics **4**(2) (2007) 131–147